STEM ACTION CENTER TECHNOLOGY PILOT

Preliminary Assessment Report February 2014-Draft for Feedback
Executive Summary

Presented by
Taylor Martin, Ph.D.
Sarah Brasiel, Ph.D.
Utah State University

LEARNING ▶ PRACTICE ▶ INSTRUCTION

Executive Summary

In the 2013 Utah Legislative session, HB 139 was passed which created the STEM Action Center for the State of Utah. This legislation also authorized a pilot program to begin in the 2013/14 school year. The pilot has two components. The first is to provide educational technology to support mathematics instruction for students in grades 6-8. The second is to prepare secondary students for college mathematics courses. The goal of this assessment is to determine the effectiveness of eight pilot programs implemented in selected schools in Utah during the 2013/14 academic year and two programs implemented in during the 2012/13 academic year. For this report we present findings from the product assessments, which are short term outcomes of this project outlined in the logic model shown in Figure 1. The long term outcomes will be measured fall 2014.

Figure 1 Logic Model for the STEM Action Center Technology Pilot

Inputs	Implementation	Short Term Outcomes	Long Term Outcomes
•Grade 7/8 mathematics software •Grade 10 college math readiness software	•1 day professional development •Implementation support from provider	•Usage of mathematics software by students •Progress in mastery or performance as measured by technology assessment •Increased interest in mathematics	•Increased mathematics achievement on the state assessment •High school mathematics readiness •College mathematics readiness

The following products were selected to include in the pilot assessment:

- SuccessMaker®, Pearson Education
- ALEKS®, ALEKS Corporation
- ST Math®, MIND Research Institute
- Think Through Math, Think Through Learning Inc.
- Math 180, Scholastic
- The STEM Academy®, The STEM Academy Inc.
- EdReady, Monterey Institute for Technology and Education
- Defined STEM, David Enterprises, Inc.

The impact of two additional programs currently implemented in Utah schools will also be evaluated based on last year's state assessment data. These programs are MATHia from Carnegie Learning and ConnectEd from McGrawHill Education.

For the assessment of the effect of each product on student mathematics performance, the best methodological approach is to use a **Cluster Randomized Control Trial design**, which is the least biased method of determining the effectiveness of a product and recommended when desiring information to inform policy or large scale implementation across a state. For every product we worked with schools to see if they would be willing to randomly assign each teacher's class sections to either intervention or control. For products where schools were unable to have random assignment, we analyze the data using a **Single Group Pre-Post Design**. We have less confidence in our results from this design, because there may be some bias since there was no control group. However we can use this type of analysis to look at relationships between the amount of usage of the product and the change in student performance from pretest to posttest, which is important to understand for future use of the products. To assess the effects of the two products used last academic year, we use a **quasi-experimental** approach for creating a comparison group using data for the entire state.

There were 46 schools that signed a memorandum of agreement and participated in the pilot. This included 118 teachers who attended professional development and used the products with their students. There were 5,722 students who used the technology products during the pilot period from October through December 2013. There was varying levels of usage within each school selected to use the product and across schools using the product. This is expected when teachers and students begin implementation of a new program under normal circumstances rather than a highly controlled laboratory experimental setting. However, if products can be shown to be effective under normal school circumstances, that is important for stakeholders and decision-makers to understand when considering purchasing products to scale up across students in a school or district.

What follows is a summary of the findings for each assessment question.

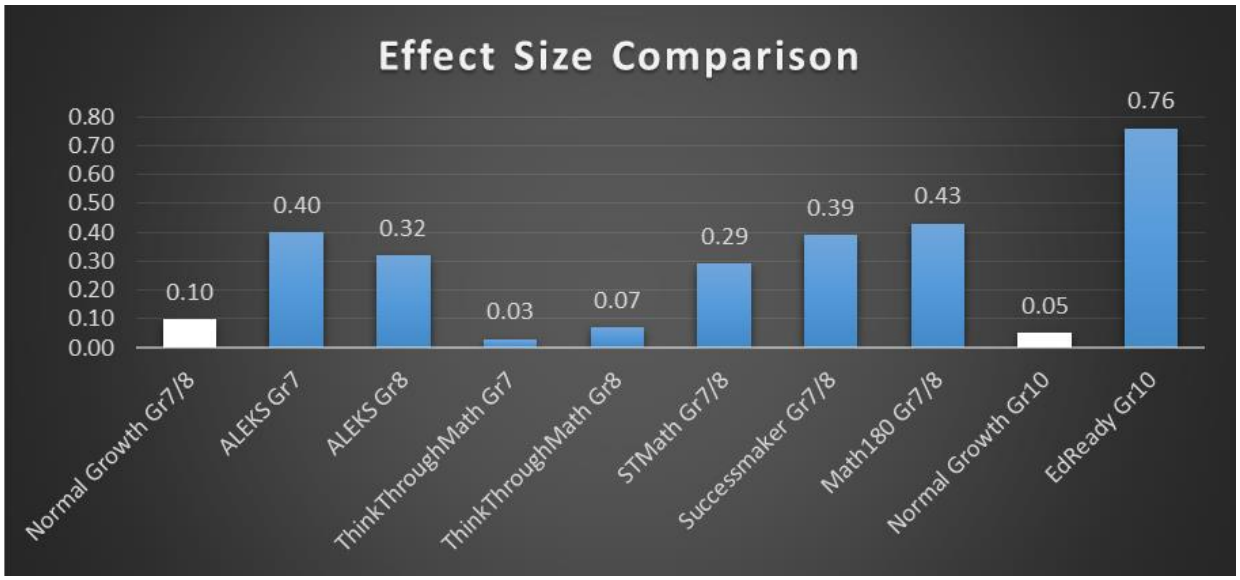
Was there a significant increase in mathematics proficiency from the start of product use (October) until the end of December in comparison to a benchmark of academic growth expected under normal circumstances in that time period?

Since each product had its own assessment with its own measurement scale, we standardized the difference between a student's pretest and posttest score, hereafter called a gain score, by creating a standardized mean difference called an *effect size*. The effect size was calculated by dividing the gain score by its standard deviation.[1] Because the effect size for each product assessment is reported in standard deviation units, the effect size for each product assessment can be compared. When just

[1] Calculation of effect size used is (gain score)/[standard deviation of the gain score/(square root of 2 times 1-r, the correlation between pretest and posttest)] from Wilson, D. B. (2001). Practical meta-analysis (Vol. 49). Sage.

looking at the gain score of students using the product it is important to compare their gain to an accepted benchmark[2] that represents the average gain score. This is expressed as an effect size that would occur under normal school year conditions and without attributing this gain to an intervention. For grade 7 and 8, a student's annual mathematics growth is an effect size of 0.32 standard deviation units, which is about 0.032 units per month. For grade 10 students, annual mathematics growth is normally 0.15; which is 0.015 per month. For the 3 month period of the pilot, we would expect a natural growth of 0.10 for grade 7/8 and 0.05 for grade 10. The graph in Figure 2 demonstrates the effect size of the gain made in student mathematics performance by product.

Figure 2 Effect Size of Gain in Mathematics Performance Compared to Natural after 3 months



In only three months, grade 7/8 students using ALEKS, Successmaker, and Math 180 made as much progress as is normally expected in an entire academic year; students using ST Math improved almost three times as would be expected. Grade 10 students using EdReady made five times the progress expected in an entire academic year with just 3 months of use. The results for Think Through Math are confounded by the fact that their outcome measure is a single digit grade level measure; therefore, it lacks the potential for variability that is needed to detect change in such a short period of time. Due to positive teacher and student responses to use of Think Through Math, we suggest waiting to draw conclusions about its effectiveness until the end of the year state assessment results are available.

[2] Annual achievement gains under normal conditions taken from page 28 of Lipsey, M.W., Puzio, K., Yun, C., Hebert, M.A., Steinka-Fry, K., Cole, M.W., Roberts, M., Anthony, K.S., Busick, M.D. (2012). Translating the Statistical Representation of the Effects of Education Interventions into More Readily Interpretable Forms. (NCSER 2013-3000). Washington, DC: National Center for Special Education Research, Institute of Education Sciences, U.S. Department of Education

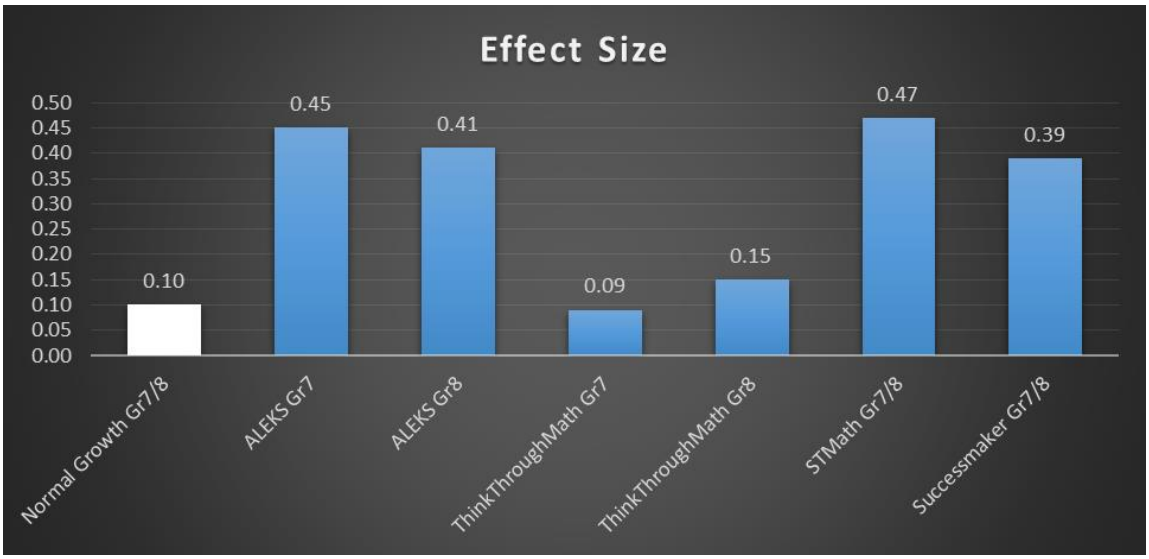
Did students in schools selected to use the technology product attain a higher level of mathematics proficiency (using the product assessment) than students in the control group (controlling for pretest proficiency)?

The gains presented in Figure 2 may not be completely related to the products, but may relate to other factors. A Cluster Randomized Control Trial design is the least biased method of determining the effectiveness of a product and is recommended when desiring information to inform policy or large scale implementation across a state. For every product we worked with schools to see if they would be willing to randomly assign teachers (or class sections) to either intervention or control. We were able to collect data for four products using a randomized control trial: ALEKS, Think Through Math, STMath, and Successmaker. Using this data, we conducted an analysis called Hierarchical Linear Modeling that accounts for students being clustered in classes in a school to determine if there was a statistically significant difference between mathematics gains made by students using the product and students not using the product (control group). ALEKS was the only product where a statistically significant difference ($p<.01$) was found where students using the product outperformed students not using the product.

Did students in schools selected to use the technology product use the product at the level recommended by the product provider?

We also conducted an analysis by reducing the sample to only students who met the benchmark set by the provider for fidelity of implementation, which is a level of implementation recommended by the provider to lead to improved outcomes for students. Although this is important to consider, it is more similar to results that occur in a laboratory school program when programs are tested under ideal circumstances rather than normal circumstances where students are absent and experience varying levels of implementation. However, we provide the effect size comparison using ideal implementation for consideration. Not all products were able to provide a benchmark for fidelity of implementation, but the ones that did provide us with that information are included in Figure 3. Using this reduced sample of students meeting the benchmark there was a slight increase in effect size for all products except Successmaker, where the effect remained the same.

Figure 3 Effect Size of Gain in Performance in 3 months for Usage that Meets Fidelity Benchmark



It is important to wait to draw conclusions about the effectiveness of the pilot products until the end of the academic year, when we will have the findings from the state assessment. This is a preferred method, since we are comparing effectiveness across products using the same outcome measure.

Did students in schools using MATHia and ConnectEd during 2012/13 perform statistically significantly higher on the state assessment in mathematics compared to similar students in the state?

There were two new products already in use in Utah, hypothesized to improve student mathematics achievement: ConnectED (McGraw-Hill) and MATHia (Carnegie Learning). We assessed the effect of these products on grade 7/8 mathematics achievement on the state assessment spring of 2013. There were 1,652 grade 7 and 1,710 grade 8 students using MATHia, 7,420 grade 7 and 7,338 grade 8 students using ConnectEd. Using data for the entire state we compared the performance of students using the products to the performance of students not using the products (36,519 grade 7 and 35,399 grade 8 students) using Hierarchical Linear Modeling with covariate adjustment. This method controlled for differences between students in prior year math achievement (baseline measure) and differences in student characteristics, school characteristics, and locale (e.g. rural). This design is the next best approach, if random assignment is not possible and when there is a small number of schools using the product. Statistically significant differences ($p<.05$) were found for ConnectED and MATHia products for grade 8; students using the products outperformed students not using the products**.** A positive difference was found for grade 7 students using these products, but the difference was not large enough to be statistically significant.

Was there an increase in student interest and engagement in mathematics for students selected to use the technology product compared to students in the control group based on pre/post survey responses?

We also analyzed data collected on changes in student engagement and interest in mathematics that may be related to use of these technology products. Research has shown a positive relationship between achievement and student engagement, self-confidence in learning mathematics, interest, motivation and self-efficacy in mathematics (Eccles and Jacobs, 1986; Helmke, 1989; Reynolds and Walberg, 1991, 1992). Therefore, we hypothesize that use of technology products may be one way to increase student beliefs about their own ability to do mathematics and their interest in pursuing a STEM career in the future. Having only experienced the pilot for 3 months we did not expect significant changes, which is why we will provide students another opportunity to take this survey at the end of the year.

The percent of students using pilot products completing the pre/post survey ranged from 3 percent for ST Math to 54 percent for Math 180. Many of the control students did not take the survey, making a comparison between students using the product and control students impossible for some products. There are some statistically significant findings from a comparison of data collected from the pre-survey in September/October 2013 and the post-survey data collected December/January 2013/14. For three products—ALEKS, Think Through Math, and ST Math—students made significant changes in how much effort they feel they have to exert to do mathematics compared to other subjects. This means students using these products are expending less effort when they do math as they felt they expended at the start of the pilot.

There were also some negative differences between pretest and posttest for students using certain products. For ALEKS, Math 180, and STEM Academy there was a decrease in student belief in their ability to do mathematics and in their perception of the value of mathematics. For Math 180 there was a decrease in student interest in mathematics, and an increase in difficulty. For STEM Academy there was also a decrease in the perception of the importance of mathematics for students. Overall there was a decrease in the overall mathematics interest and engagement scores from pre-survey to post-survey for Math 180 and STEM Academy.

One explanation for the drop in scores, could be that many students took the survey prior to winter break, when mathematics may be of less interest as they think about having time off school. Another explanation to consider is whether initial use of these products brings to student's attention that they do have gaps in their knowledge. Although the product is set up for students to succeed, they may initially pay more attention to the large amount of content they still need to master, as shown by graphs or progress monitors within the product that allow student to assess their own growth.

Were teachers who were selected (or had class sections selected) to use the technology product with their students satisfied with the product and student response/outcomes?

Finally, we summarized teacher feedback from monthly surveys by product to understand the satisfaction and concerns they have with each product. When teachers were asked how they used the product in their classroom, a majority (40 percent) of the teachers described using the product with students for grade level instruction. Thirty-five percent responded that they are using the product for below grade level instruction, and twenty-nine percent are using the product for above grade level instruction. Only nineteen percent use the product for homework. In fact, only three products were used for homework: ALEKS, Think Through Math, and STMath. The most common positive teacher response was related to the level of engagement of their students, which included responses for ALEKS, Think Through Math, and STMath. The second most common response was the way the program was individualized to student needs, which included responses from teachers using those same three products in addition to the EdReady product. It is important to note that one teacher responded that ST Math was being used with success with Honors students, since this product was viewed by many teachers as a product more for remediation. There was no positive response of satisfaction from any teacher using Successmaker or Defined STEM.

The greatest concern or area of feedback for improvement was teachers expressing the need for additional training for use of the product. The second greatest concern was with the lack of alignment of the product to the standards the teachers are expected to teach; this was limited to two products (Successmaker and STEM Academy). Some of the teachers seem to prefer high levels of teacher control wanting to select the content for students, wanting to design their own problems for students, and set the path within the content for students to take. However, some of the products purposely allow student choice, as students take responsibility for their own learning. This is more of a philosophical difference in what is a more effective instructional approach. There were also concerns related to specific technology products, such as complaints from teachers about products requiring Flash® that needs to be updated periodically or outdated computers that are slow when running the programs. Other concerns had to do with availability of technology, where teachers complained of having to sign up for a computer lab and not always having access to computers for their students.

There are several limitations to consider. There are limitations associated with the use of mathematics performance measures within each product for this pilot, since each measure was a different type. It is important to wait until the end of the year when the state assessment results are available, since that will be a valid measure of mathematics achievement across all products. There is also a limitation in the mathematics interest survey findings in that there was a small sample size of students completing both a pre-survey and post-survey, with most products having less than 50 percent of students responding.

There was also a limitations associated with use of Teacher Survey data, which is a self-report method of data collection. The data could be inaccurate or slightly inflated since teachers know they are participating in a pilot, and may want to appear to be following expectation to participate in the pilot. It is also important to understand that the reason there is a diversity of sample sizes by product for this pilot, is because the size was limited by the available donations by product providers.

This pilot would not have been possible without the generous donation of close to half a million dollars in software, licenses, and implementation support from the product providers.

Product	Product Provider	Donated Licenses, Memberships, or Materials	Approximate Value of Donation
ALEKS®,	ALEKS Corporation	1,739 students in 11 schools	$103,462.50
ST Math®	MIND Research Institute	1,596 students in 8 schools	$153,000.00
Think Through Math	Think Through Learning Inc.	2,152 students in 7 schools	$49,597.00
Successmaker	Pearson Education	765 students in 4 schools	$49,500
Math 180	Scholastic	Already in use with 82 students in 3 schools	$0.00
The STEM Academy®	The STEM Academy Inc.	14 schools (data not collected at student level)	$135,945.00
EdReady	Monterey Institute for Technology and Education	206 students in 5 schools	$3,000.00
		Total Value of Donations	$494,504.50

Next we will be analyzing back-end data files from the providers in order to provide additional information about student use of the products. For example, mathematics content information can help us understand the types of mathematics content students have been exposed to using the technology, which will help us understand if they have been experiencing learning opportunities related to topics they will be tested on in the state assessment. We have received back-end data from ST Math and EdReady. We will make another request to the other products and then conduct learning analytics to understand usage of these products in terms of mathematics topics to provide context for the state assessment results. These analyses can also take into consideration student demographic differences.

In October 2014 the spring 2014 state assessment results should be available for analysis. We will compare the performance of students in pilot schools using the products to students in the state who did not have the opportunity to use these products to determine their effectiveness in improving

outcomes for students. We can also disaggregate the data to look at performance of students by different demographic group to see if the achievement gap is closing for students using certain products.

Once we have the effect sizes for each product on improving student achievement on the state assessment, we will conduct a cost-effectiveness analysis to determine— given the cost of the product and the effect size—which product is the most cost-effective. This type of analysis is referred to as a *return on investment* analysis. It can be used to inform future purchasing decisions. When resources are limited, as they often are in education, finding products that produce the largest effects for a given per student cost can have the greatest practical significance for stakeholders who need to make purchasing decisions.

Contact Information

TAYLOR MARTIN
DIRECTOR

taylormartin@usu.edu

SARAH BRASIEL
ASSOCIATE DIRECTOR

Sarah.brasiel@usu.edu

UtahStateUniversity

Active Learning Lab

LEARNING ▶ PRACTICE ▶ INSTRUCTION